Who Is A Student:

Completion in Coursera Courses at Duke University

Molly Goldwasser, Chris Mankoff, Kim Manturuk, Lorrie Schmid, Keith Whitfield[1]

Duke University

Molly Goldwasser

(phone) 919.684.0731

(fax) 919.684.4421

(email) molly.goldwasser@duke.edu

[1] Molly Goldwasser, Trent Hall Office 007, 310 Trent Drive, Durham, NC 27710
Chris Mankoff, Trent Hall Office 008, 310 Trent Drive, Durham, NC 27710
Kim Manturuk, 030 Bostock Library, Duke University, Durham, NC 27708
Lorrie Schmid, Gross Hall Room 245, Durham, NC 27705
Keith Whitfield, 220 Allen Building, P.O. Box 90005, Duke University, Durham, NC 27708

Abstract

Much of the interest in MOOCs centers on questions about who completes these sorts of courses. Duke's Coursera-based Massive Open Online Courses (MOOCs) confirm many demographic trends previously delineated by researchers at peer institutions. As found in previous research, this study found individuals who speak English as a first language and who already earned at least a Bachelor's degree are the most likely to complete a Coursera course. MOOC researchers to date have not, however, developed clear operational definitions about who constitutes a learner in these online courses. This paper proposes some possible definitions to standardize future research. Further, this study looked at factors that predict different learner participation levels and investigated which activities predict Coursera course completion. Study results indicated that viewing online forums and participation in online discussions are both predictive of course completion. The findings suggest that more than demographics are important in understanding learner behavior in this alternate delivery of knowledge; participation in online discussions and viewing discussion forums are both important factors to consider.

Key words: Massive Open Online Courses (MOOCs), Coursera, Completion, Duke University

Introduction

Online and distance education– Massive Open Online Courses (MOOCs) in particular – have received a lot of publicity in the last couple of years. One systematic literature review identified eight key areas of MOOC research study. They are: introducing the concept of the MOOC, challenges and opportunities of MOOCs in higher education, examining a single course or a set of courses, examining pedagogic approaches, examining hardware/software solutions, issues of learner participation, issues of provider participation, and other topics not covered in the following areas (Liyanagunawardena, Adams, & Williams, 2013). MOOCs are typically virtual, distributed classrooms that exist for an average of six to ten weeks at a time, including instructional videos, regular assessments, and communication forums, all within a single, integrated computer platform.

Originally heralded as a platform to increase access to, and affordability of, educational content, challenges arise as professors scale up their courses to reach massive audiences (Pappano, 2012). Within higher education discourse, a key criticism of MOOC participation has been the low completion rate among its learners (Kolwich, 2013). With well over 10,000 enrollments at the outset of most courses, completion rates (as percentages of original enrollment numbers) are low (Catropa, 2013; Jordan, 2014). Kolowich (2013) suggests the overall completion rate of MOOCs hovers around ten percent. While the percent of learners who complete MOOCs is low, the number of completers is not insignificant. As of September 2014, Duke University had issued over 55,000 Statements of Accomplishment and Verified Certificates (“MOOCs at Duke”, 2014). This report examines how learner characteristics and activities impact Coursera course completion.

Most MOOC research has focused on a single course or a set of courses within the same discipline (e.g., Kizilcec, Piech, & Schneider, 2013). Research on learner participation has focused on aspects that can be altered within a course, including peer assessment and emotive vocabulary

(e.g., Koutropolous et al., 2010; Kulkarni et al., 2013). Examinations of collections of courses will allow effects across types of courses over different subjects and disciplines generalize (or not) to learner effects observed in these courses.

While little has been published on learner engagement and motivation, and the effects of those on performance in MOOCs, before addressing these issues researchers need to examine a critical issue yet to be resolved in the MOOC literature: who is a student?

Who Is A Student?

MOOC enrollment and persistence statistics consistently classify completers as those who have earned some form of a certificate of achievement (in Coursera, these include a Statement of Accomplishment or a Verified Certificate) from the MOOC provider. There is no consensus about who constitutes a student at the front end of the course. Is a student someone who:

- Enrolls in the course?
- Visits the course website?
- Watches a course video?
- Completes an assignment?
- Participates in a discussion forum?

Traditional education systems wait until the end of a grace period (such as a “drop/add” period) to count enrollment and to determine baseline student statistics. If MOOC researchers were to do the same, course completion statistics would increase.

Across the 18 course offerings at Duke between 2012 and 2014, nearly 900,000 learners enrolled. However, as we began looking at the data, it was clear that many, if not most, of the learners did

not complete many of the key components of the course. Therefore, the total enrolled numbers may be artificially inflated with those who enrolled but never interacted with any part of the platform. In addition, useful information about when and how individuals use the system, regardless of their completion status, can be informative to understanding learner engagement with the system (Kizilcec et al., 2013).

As research in the distance and online education field continues, it is vital that the research community consider a standard definition of what constitutes a MOOC student. The authors of this article posit that the larger issue in the MOOC discourse is not one of low retention but rather a lack of clarity regarding who constitutes the student body.

Methods

This set of analysis included all enrolled learners in 18 unique course session offerings, which comprise 58% of the offerings given at Duke between 2012 and 2014. All courses with complete data were included, the dropped courses were due to data problems found in the sources files for the course. See Appendix 1, Table 1 for enrollment and activity counts by course. Data was collected in two main ways: through the Coursera platform and through the use of a pre- and post-survey designed by the Center for Instructional Technology (CIT) at Duke. The Coursera data includes demographic data to be used in the analyses as predictors and/or control variables. Demographic data include: age, gender, educational level, English as a primary language, race, ethnicity, nationality and employment status. Learners can view and post on a Coursera online forum. Finally, the Coursera platform tracks how each individual interacts within the online classroom through their use of clickstream technology. Learner involvement across these data platforms and activities are found in Table 2. All data is stored and analyzed on a protected research data network (PRDN), created especially to handle sensitive and restricted data.

Approximately 900,000 learners enrolled in these 18 course session offerings. Fifty-five percent of the learners identified as male, and 45% identified as female. Learners from all over the world and many nationalities are represented and demographics on race/ ethnicity reflect that. Sixty-three percent identify as White, 22% as Asian, 4% as Black, and 8% as some other category. Sixty-two percent of the sample was aged 34 and younger. Across the whole sample, 35% completed a Bachelors and an additional 30% had advanced degrees. Forty-eight percent work full time.

Two sets of interconnected logistic regression analyses were conducted, one examining which course activities were completed and one focused on identifying which activities predicted course completion for learners with different demographic profiles. Course activities included: visiting the course website, watching one or more course videos, viewing or writing a forum post, completing one or more assignments, and receiving a certificate of completion. A set of dummy variables were created for these analyses. For example, the male dummy variable sts all male learners to "1" and female learners to "0". The reference group described in the intercept is all female learners and the male dummy variable describes the differences between male learners and the intercept. Dummy variables were constructed for gender, race, English as a first language, educational status, and age. The reference groups are male, white, English as a second language learners, college graduates, and aged 18-24.

Results: Duke Data Confirms Trends

Recent research from Harvard and MIT suggest several trends in MOOCs (Newman & Oh, 2014):

- Course completion rates are low if you begin with who enrolled in the course.
- Most completers are college graduates.

- More men enroll in courses than women.

Data from Duke's Coursera courses also follow these trends. As described above, reported course completion rates vary depending on how the researcher defines a student. Using the data defined within the methods, of those who complete at least one assignment (n= 192,682), 13% ultimately complete the course (complete all assignments and receive a final grade). Even this elevated statistic is low compared to course completion rates among students in brick and mortar courses. Eighty percent of the learners who complete Duke MOOCs already have college degrees and 52% of the course completers are men.

Participants Throughout the Course

A series of logistic regression models were conducted to determine which types of learners were more likely to complete which activities (see Tables 3 and 4). Learner characteristics for each model included: gender, race, English as a first language, educational status, and age. The first regression compared whether or not someone who enrolled in the course watched at least one video. Hispanic / Latinos learners were slightly more likely to watch a video than other races (Odds Ratio *OR* = 1.18). In addition, older learners were more likely to watch a video compared to younger learners (*OR* for learners 65 and older, = 2.38). The next regression compared whether or not a learner viewed the forums. African American learners were less likely to view the forums compared to European American learners (*OR* = 0.73). Learners where English was their first language were slightly more likely to view the forums (*OR* = 1.12). Again, the trend of older learners being more active than younger learners continued, with older learners more likely to view the forums (*OR* for learners 65 and older = 2.07). Slightly different trends were found for learners who wrote a forum post. Although English as a first language (*OR* = 1.21) and age (*OR* for age 65+ = 1.61) continue to be important, significant results are also found for gender and education.

Specifically, males are less likely to write a forum post than females (*OR* = 0.78), and those with more than a Bachelors' degree are less likely to write a forum post than those with a Bachelors degree (*OR* = 0.85).

Having explored the factors that predict different levels of course participation, we return to the question of who is a student. To address this point, we analyzed whether the different learner participation types we identified are significant predictors of course completion.

Completers

Logistic regression models using Duke and Coursera data indicate several significant findings. Learners who are most likely to complete Duke MOOCs:

- Speak English as a first language (*OR* = 1.25).
- Earned a degree higher than a Bachelor's degree (*OR* = 1.22).
- Are between the ages of 35-64.

Another key finding is that learner activity impacts course completion. Participation in online forums and online discussions matter; participants in these activities were more likely to persist to course completion. While participation impacts completion in both statistically and substantively significant ways, the exact amount of time spent in online forums and online discussions does not. The extent to which a learner participates in online activities was not statistically significant with regards to predicting course completion.

Discussion and Next Steps

The findings of the current study highlight the importance of defining what a student is when looking at patterns of participation in MOOCs. The results here are similar to the few other studies currently reporting on completion rates. For learners enrolled in Duke Coursera courses, viewing online forums and participation in online discussions are both predictive of course completion. While the effects are small, there is a positive correlation between both the frequency of online forum views and frequency of participation in online discussions and course completion. There remains a need to standardize the definition of a student in the MOOC literature. We propose that participants who complete at least one assignment define this population.

Important in these findings is that education, age and gender matter in distinctive ways depending on how one defines the population of interest. Our results suggest that older learners, while a smaller proportion of the overall population of MOOC learners, is more likely to watch a video but less likely to complete the course than younger participants. These differences may indicate that learners of different ages may have different intentions when registering for a MOOC.

The effects that were found for gender are very interesting. While more men enrolled than women, women were more likely to engage with the course by writing a forum post. Many instructors of MOOCs are interested in the utility of the forums for discussing course material and creating community among geographically diverse course participants. Our results indicate that, while most learners do not participate in the discussion forums, those who do are more likely to complete the course.

Demographic variables in this study were defined by traditional American classifications. Additional research is needed to examine student trends by sub-category according to different global norms. Future research is also needed to investigate how lessons learned from Duke professors' MOOCs can impact traditional students on campus. For example, researchers can look at participation as a

proxy for engagement in brick and mortar classes on campus – how do students who post in online forums and online discussions fare compare to those who do not? Are these findings similar to those in online courses?

Acknowledgments

Thank you to Dr. Carol Ripple, Associate Director for Education Research and Engagement, and Dr. Lynne O'Brien, Associate Vice Provost for Digital and Online Initiatives for their support of this study.

References

Catropa, Dayna (24 February 2013). "Big (MOOC) Data". *Inside Higher Ed*. Retrieved 22 September 2014.

Jordan, Katy. "MOOC Completion Rates: The Data". Retrieved 22 September 2014.

Kolowich, Steve (8 April 2013). "Coursera Takes a Nuanced View of MOOC Dropout Rates". *The Chronicle of Higher Education*. Retrieved 22 September 2014.

Kizilcec, R. F., Piech, C., & Schneider, E. (2013). Desconstructing disengagement: Analyzing learner subpopulations in massive open online courses. In *Proceedings of the 3rd International Conference on Learning Analytics and Knowledge* (pp. 170-179).

Kulkarni, C., Wei, K. P., Le, H., Chia, D., Papadopoulos, K., Cheng, J., Koller, D., & Klemmer, S. R. (2013). Peer and self assessment in massive online classes. *ACM Transactions on Computer-Human Interaction, 20*(6).

Liyanagunawardena, T. R., Adams, A. A., & Williams, S. A. MOOCs: A Systematic Study of the Published Literature 2008-2012. *Review of Research in Open and Distance Learning, 14*(3), 202-227.

Massive Open Online Courses at Duke (2 September 2014). Retrieved 24 September 2014 at http://online.duke.edu/about/moocs/

Newman, Jonah and Soo Oh (13 June 2014). "8 Things You Should Know About MOOCs". *The Chronicle of Higher Education*. Retrieved 22 September 2014.

Pappano, Laura. "The Year of the MOOC". The New York Times. Retrieved 22 September 2014.